# Rational Design of Macrolides by Virtual Screening of Combinatorial Libraries Generated through in Silico Manipulation of Polyketide Synthases

Sergey B. Zotchev, Alla V. Stepanchikova, Anastasia P. Sergeyko, Boris N. Sobolev, Dmitrii A. Filimonov and Vladimir V. Poroikov

Department of Biotechnology, Norwegian University of Science and Technology

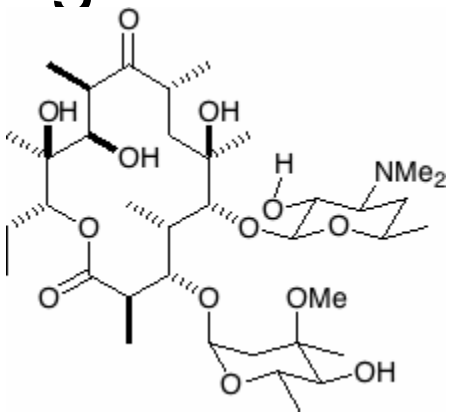
Journal of Medicinal Chemistry, 2006, 49, 2077-2087

# Outline

- Introduction
- Program Approach
- PKS Review
- Software
- Examples

# Introduction

- Natural therapeutic products
  - Random screening
- Limitations
  - Focusing on particular targets
  - Low ratio of active to inactive molecules
- Already have library building programs
  - Based on chemoinformatics
  - Target- and Ligand-based



# Program Approach

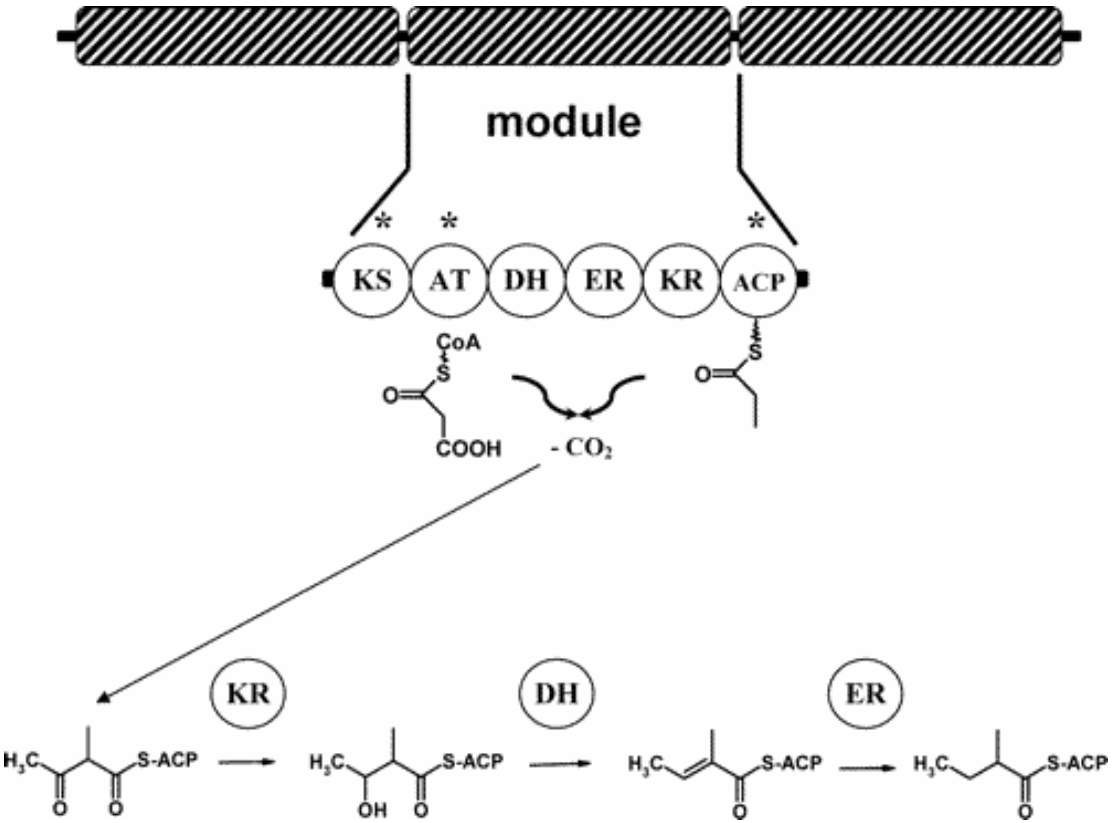
- Based on postgenomic technology and methods of chemoinformatics
- Generates secondary metabolites' structures in correspondence with the appropriate biosynthetic pathways
- Prediction of properties (biological activity, physiochemical properties, etc)

# Approach Testing

- Tested on macrolides
- All combinatorial manipulations:
  - Module and domain insertions/deletions
  - Site specific mutagenesis
- Can screen for products with desired therapeutic properties and rule out toxic products

# Review: Biosynthesis of Macrolides by Polyketide Synthases

- Polyketide synthase type I enzymes make polyketide chains
- Chains are formed from decarboxylative condensations of carboxylic acids
- Starter units are activated carboxylic acids
- Macrolactone rings formed by cyclization of polyketide chains




KS: β- ketoacylsynthase

AT: acyltransferase

KR: ketoreductase

DH: dehydratase

ER: enoyl reductase

ACP: acyl carrier protein

TE: thioesterase

Figure 1, J. Med. Chem, 2006, Vol. 49, No 6

# Successful Manipulations

- Deletions of PKS modules
- Exchanges of AT domains
- Inactivation and/or addition of reductive domains

# Problems

- Actually performing these changes in a lab can lead to
  - Loss of functionality
  - Loss of or no product yield
  - Loss of time
  - Loss of money
- Domain changes only affect the macrolactone ring which is not the final product
- Making every change is impossible

# Biogenerator Software

- Performs in silico manipulation of theoretical PKS type I systems
- Creates virtual libraries of macrolide analogs
- Introduces post-PKS modifications
- Virtual screenings for characteristics

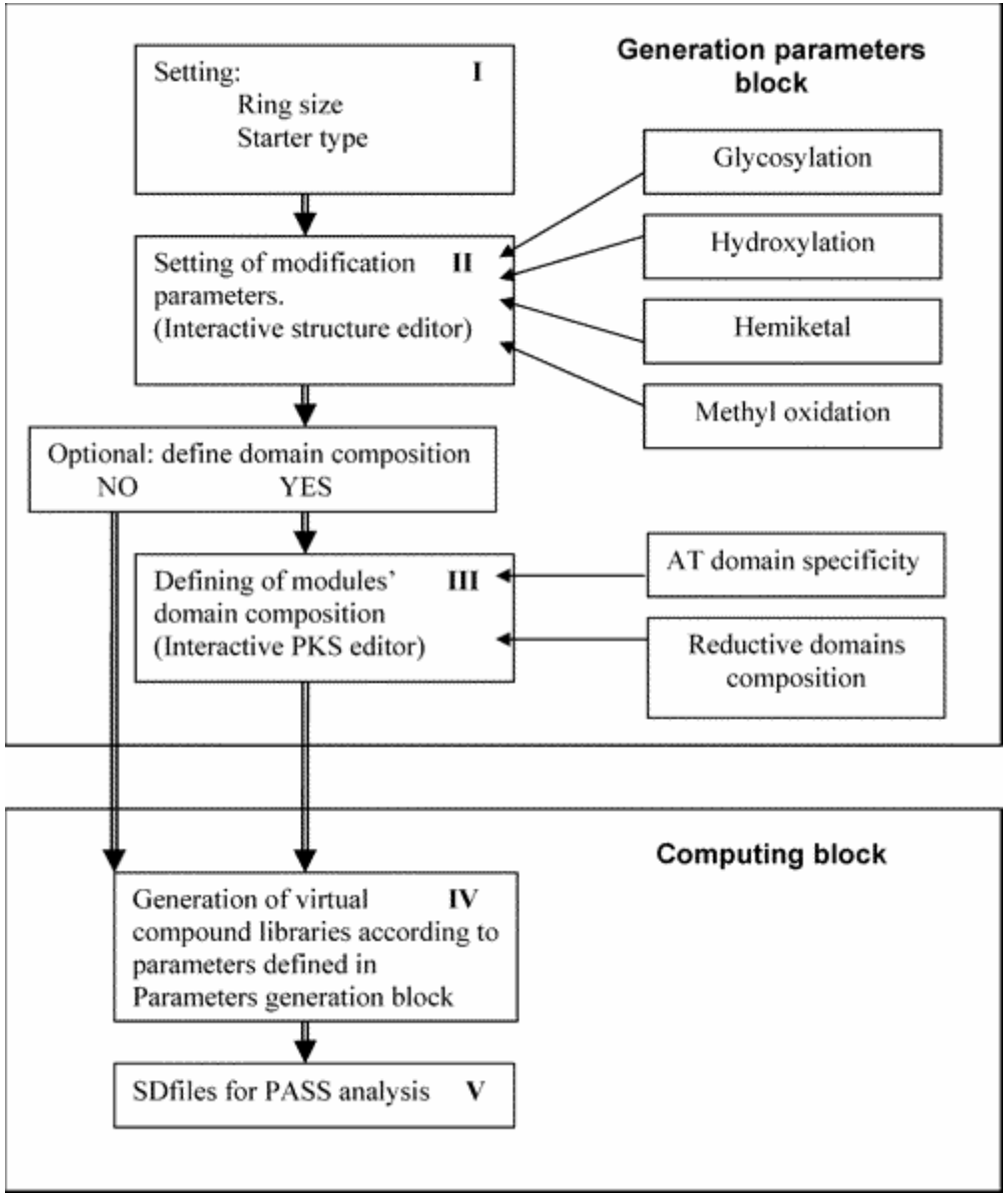



Figure 2, J. Med. Chem, 2006, Vol. 49, No 6

# Blocks

- Block I
  - Starter Type: Acetate or propionate
  - Ring Size: Up to 38 atoms (module number)
- Block II
  - Post-PKS Modification Parameters:
    - Glycosylation
    - Hydroxylation
    - Methyl Oxidation
    - Hemiketal

# Blocks

- Glycosylation: Set glycosylation sites at the putative hydroxyl groups, choose type of sugar
- Hydroxylation: Sites can be set based on knowledge of specificity of enzymes
- Methyl Oxidation: Choose if a carboxyl group is at a certain position
- Hemiketal: Restricts domain composition in two modules

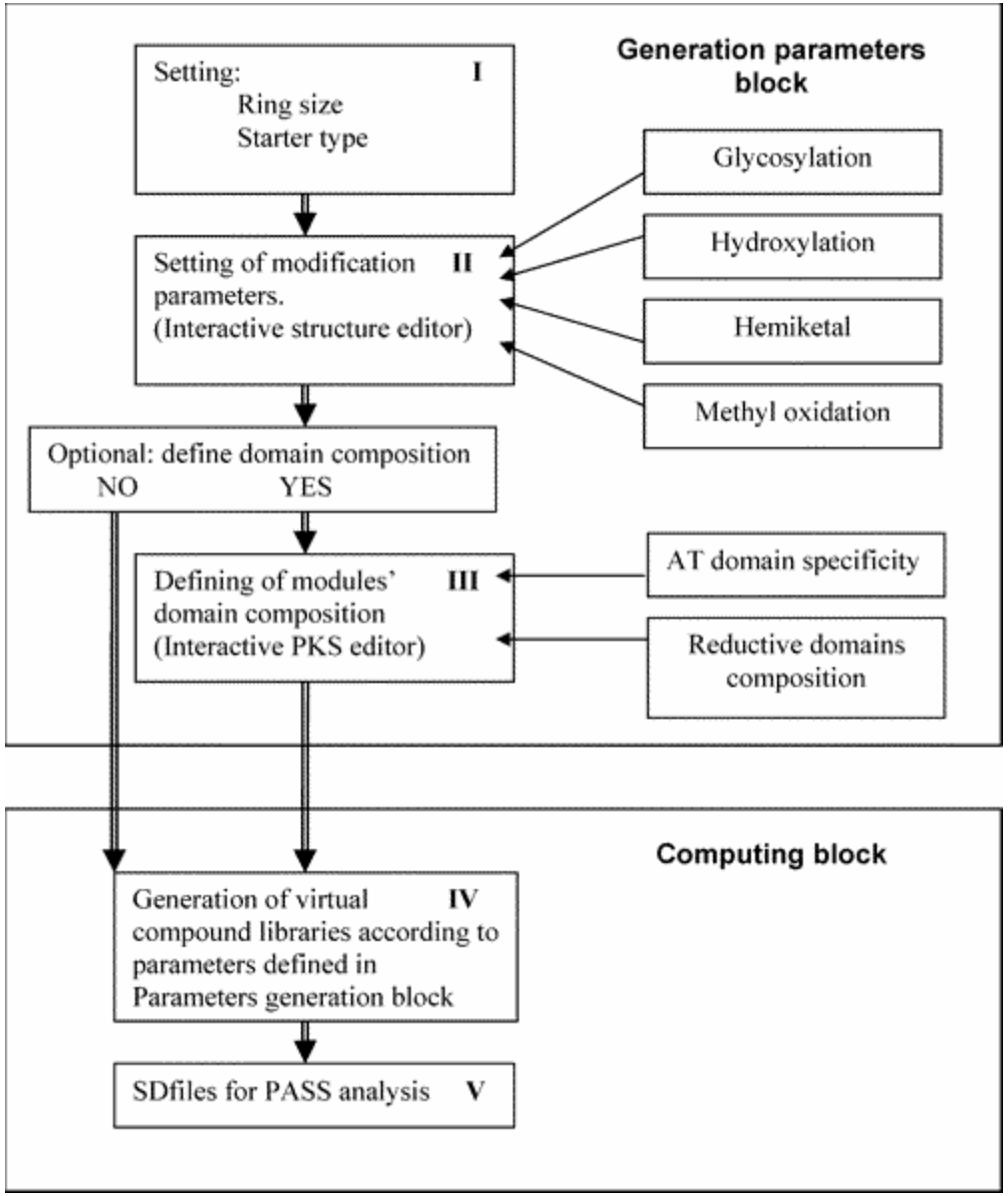



Figure 2, J. Med. Chem, 2006, Vol. 49, No 6

# Blocks

- Option to Define Domain Composition
- Block III
  - Module's Domain Composition: AT domain specificity and reductive domain composition
- Block IV
  - Generation of Virtual Combinatorial Libraries: Chemical structures and PKS notation

# Biological Activity Prediction

- PASS Software:
- 85% accuracy
- Can predict more than 2000 pharmacological effects, mechanisms of action and toxic effects based on structural formulas
- Training set of about 58,500 known substances

# Biological Activity

- Probability (Values from 0.000 to 1.000)
  - Pa: Probability to be active
  - Pi:  Probability to be inactive
- User defines threshold (T) for activities with Pa > T to be considered, default Pa=Pi
- PharmaExpert: Helps in more specific selection

# Example 1: Erythromycin

- Chosen because it has already been very well studied (control)
- Tested PASS system on 285 erythromycin analogs with experimentally confirmed activities

| Type of activity | *N* | ± | Pa | Pi |
|---|---|---|---|---|
| Acne therapy | 7 | 0 | 0.738 | 0.005 |
| Antibacterial Drugs | 1 | 0 | 0.928 | 0.004 |
| Antibiotics | 116 | 0 | 0.983 | 0.002 |
| Anti-helicobacter pylori agents | 1 | 0 | 0.202 | 0.091 |
| Antimycobacterial agents | 13 | 0 | 0.516 | 0.014 |
| Endometriosis therapy | 7 | 7 | | |
| GnRH (LHRH) antagonists | 18 | 0 | 0.587 | 0.002 |
| Macrolides | 256 | 0 | 0.993 | 0.001 |
| Motilin receptor agonists | 11 | 0 | 0.343 | 0.001 |
| NF-(kappa)B activation inhibitors | 1 | 1 | | |
| Oncolytic drugs | 11 | 0 | 0.427 | 0.095 |
| Prokinetic agents | 11 | 0 | 0.858 | 0.002 |
| Prostate cancer therapy | 7 | 0 | 0.584 | 0.003 |
| Quinolones | 9 | 9 | | |
| Treatment of protozoal diseases | 1 | 0 | 0.565 | 0.012 |
| Treatment of tuberculosis | 1 | 1 | | |
| Sum | 471 | 18 | 96% | |

Table 1. PASS Prediction of Known Activity Types for 285 Erythromycin Analogues from the MDDR Database[a]

# Combinatorial Biosynthesis

- Parameters:
  - Propionate starter
  - 6 PKS modules
  - 2 glycosylation positions
  - 2 hydroxylation positions

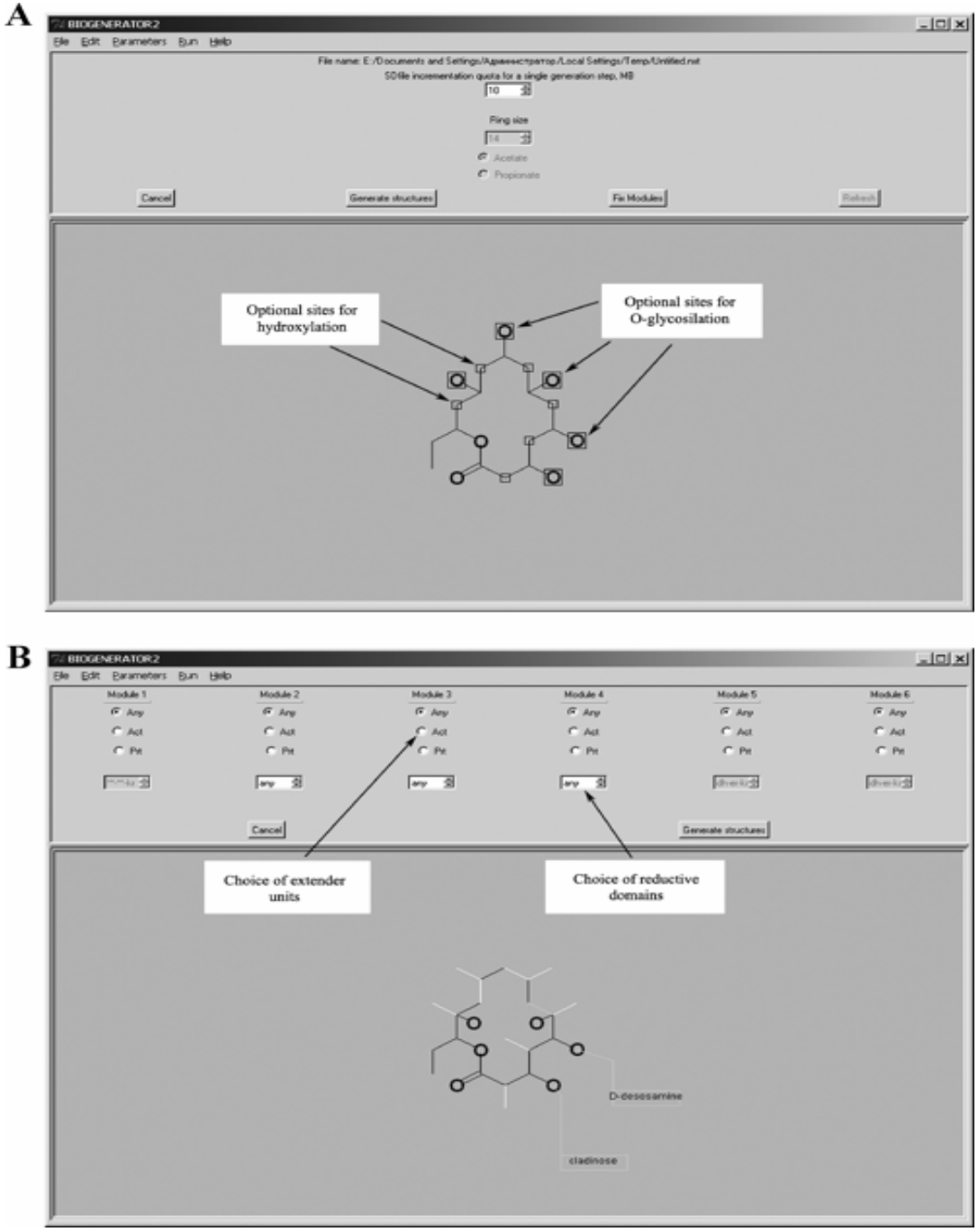
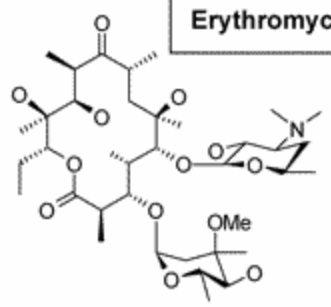



Figure 4, J. Med. Chem, 2006, Vol. 49, No 6

# Results

- 3072 structures
- Pa > 0.7
- Also analyzed 1060 known structures more structurally diverse due to no limitations on additional ring groups

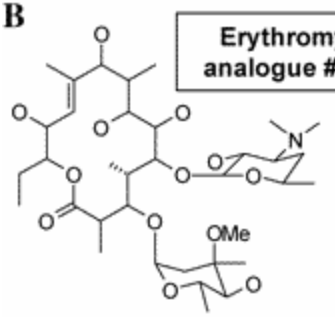
**A**




**PASS ACTIVITY SPECTRUM** at Pa > 0.7

| Pa | Pi | |
|---|---|---|
| 0.975 | 0.003 | Antibacterial |
| **0.970** | **0.004** | **Protein synthesis inhibitor** |
| 0.966 | 0.002 | Protein 50S ribosomal subunit inhibitor |
| 0.952 | 0.004 | HERG channel antagonist |
| 0.864 | 0.007 | CYP3A1 substrate |
| **0.854** | **0.008** | **QT interval prolongation** |
| 0.829 | 0.007 | CYP3A substrate |
| 0.811 | 0.006 | Antidyskinetic |
| 0.783 | 0.006 | Selectin antagonist |
| 0.766 | 0.011 | CYP3A4 substrate |
| 0.743 | 0.010 | Antiprotozoal (Toxoplasma) |
| 0.730 | 0.001 | Testosterone antagonist |

**B**




**[Loading: Propionate]**
**[KS-ATmal-**-**-KR-ACP]**
**[KS-ATmet-DH-**-KR-ACP]**
**[KS-ATmet-**-**-KR-ACP]**
**[KS-ATmal-**-**-KR-ACP]**
**[KS-ATmal-**-**-KR-ACP]**
**[KS-ATmet-**-**-KR-ACP]**

**PASS ACTIVITY SPECTRUM** at Pa > 0.7

| Pa | Pi | |
|---|---|---|
| 0.867 | 0.006 | HERG channel antagonist |
| **0.848** | **0.005** | **Protein synthesis inhibitor** |
| 0.800 | 0.004 | Antimetastatic |
| 0.798 | 0.004 | Antibacterial |
| 0.787 | 0.009 | CYP3A1 substrate |
| 0.758 | 0.021 | Cholesterol synthesis inhibitor |
| 0.746 | 0.010 | Antidyskinetic |
| 0.741 | 0.006 | Selectin antagonist |
| 0.743 | 0.012 | CYP3A substrate |
| **0.562** | **0.056** | **QT interval prolongation** |

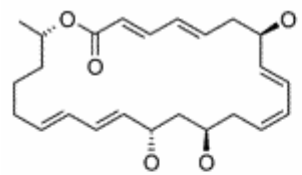
Figure 3, J. Med. Chem, 2006, Vol. 49, No 6

# Example 2: Macrolactin A

- Macrolides with unknown marine origin
- Possible activities: antibacterial, antitumor, shown to inhibit replication of herpes simplex virus and HIV, shown to inhibit cholesterol biosynthesis

# Combinatorial Biosynthesis

- Parameters:
  - Acyl-coA starters
  - Acetate extenders
- Results:
  - 1,048,576 structured library
  - Results similar to known compounds
  - Also introduced new activities



**Macrolactin A PKS:**
**[Loading: Acetate]**

**M1: [KS-ATmal-**-**-KR-ACP]**
**M2: [KS-ATmal-DH-ER-KR-ACP]**
**M3: [KS-ATmal-DH-**-KR-ACP]**
**M4: [KS-ATmal-DH-**-KR-ACP]**
**M5: [KS-ATmal-**-**-KR-ACP]**
**M6: [KS-ATmet-**-**-KR-ACP]**
**M7: [KS-ATmet-DH-**-KR-ACP]**
**M8: [KS-ATmet-DH-**-KR-ACP]**
**M9: [KS-ATmet-**-**-KR-ACP]**
**M10: [KS-ATmet-DH-**-KR-ACP]**
**M11: [KS-ATmet-DH-**-KR-ACP]**

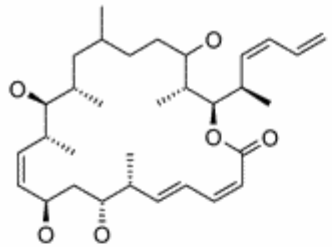
**PASS ACTIVITY SPECTRUM** at Pa > 0.7

| Pa | Pi | |
|---|---|---|
| 0.938 | 0.003 | Cholesterol synthesis inhibitor |
| 0.916 | 0.002 | Antimetastatic |
| 0.784 | 0.008 | H+-transporting two-sector ATPase inhibitor |
| 0.818 | 0.054 | Dextranase inhibitor |
| 0.737 | 0.028 | Antineolpastic (lymphocytic leukemia) |
| 0.737 | 0.028 | Antineolpastic (non-Hodgkin's lymphoma) |
| 0.759 | 0.059 | Phosphatase inhibitor |
| 0.711 | 0.012 | Apoptosis agonist |
| 0.709 | 0.096 | Acylcarnitine hydrolase inhibitor |
| 0.703 | 0.009 | CYP2A11 substrate |

Figure 5, J. Med. Chem, 2006, Vol. 49, No 6

# Results

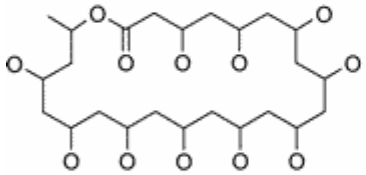
- Macrolactins structurally resemble dictyostatin-1, known to have tubulin stabilizing activity (mechanism of antitumor agents)
- 236 of the macrolactin analogs had microtubule stabilizing activity, Pa> 0.7



**Dictyostatin-1**

**PASS ACTIVITY SPECTRUM** at Pa > 0.7

| Pa | Pi | |
|---|---|---|
| 0.922 | 0.006 | Phosphatase inhibitor |
| 0.895 | 0.005 | Cholesterol synthesis inhibitor |
| 0.827 | 0.006 | H+transporting two -sector ATPase inhibitor |
| 0.803 | 0.004 | Antimetastatic |
| **0.784** | **0.003** | **Microtubule stabilizator** |
| **0.744** | **0.005** | **Microtubule formation stimulant** |
| 0.727 | 0.008 | Immunosuppressant |
| 0.723 | 0.004 | Membrane Permeability enhancer |
| 0.762 | 0.044 | Beta-adrenergic-receptor kinase inhibitor |
| 0.715 | 0.011 | CYP2A substrate |



**Marolactin analogue # 49526**

**PASS ACTIVITY SPECTRUM** at Pa > 0.7

| Pa | Pi | |
|---|---|---|
| 0.951 | 0.003 | Cholesterol synthesis inhibitor |
| 0.929 | 0.004 | Antineolpastic (lymphocytic leukemia) |
| 0.929 | 0.004 | Antineolpastic (non -Hodgkin's lymphoma) |
| 0.910 | 0.016 | Dextranase inhibitor |
| 0.863 | 0.003 | Antimetastatic |
| 0.845 | 0.040 | Acylcarnitine hydrolase Inhibitor |
| 0.778 | 0.052 | Flavanone 4-reductase inhibitor |
| 0.773 | 0.015 | Membrane integrity Antagonist |
| 0.771 | 0.054 | Phosphatase inhibitor |
| 0.762 | 0.007 | Apoptosis agonist |
| 0.759 | 0.006 | Pectin lyase inhibitor |
| **0.749** | **0.003** | **Microtubule stabilizator** |
| 0.755 | 0.012 | Bcl2 antagonist |
| 0.745 | 0.007 | CYP3A5 substrate |
| 0.747 | 0.016 | Glycerol-1-phosphatase inhibitor |
| 0.736 | 0.010 | Atherosclerosis treatment |

Figure 6, J. Med. Chem, 2006, Vol. 49, No 6

# Conclusions

- Results of the software give “instructions” for molecular biologists to follow in the construction of desired products
- Reasonable accuracy
- Currently experimentally testing some samples